Generalized Symmetric Divergence Measures and the Probability of Error

Inder Jeet Taneja

Departamento de Matemática
Universidade Federal de Santa Catarina
88.040-900 Florianópolis, SC, Brazil.
e-mail: taneja@mtm.ufsc.br
http://www.mtm.ufsc.br/~taneja

Abstract

There are three classical divergence measures exist in the literature on information theory and statistics. These are namely, Jeffryes-Kullback-Leiber [6], [8] J-divergence. Sibson-Burbea-Rao [9], [3] Jensen-Shannon divegernce and Taneja [11] Arithmetic-Geometric divergence. These three measures bear an interesting relationship among each other. The divergence measures like Hellinger [5] discrimination, symmetric χ^2- divergence, and triangular discrimination are also known in the literature. In this paper, we have considered generalized symmetric divergence measures having the measures given above as particular cases. Bounds on the probability of error are obtained in terms of generalized symmetric divergence measures. Study of bounds on probability of error is extended for the difference of divergence measures.

Key words: *J-divergence; Jensen-Shannon divergence; Arithmetic-Geometric divergence; Probability of Error, $f-$divergence*

AMS Classification: 94A17; 62B10.

1 Introduction

Let

$$\Gamma_n = \left\{ P = (p_1, p_2, ..., p_n) \left| p_i > 0, \sum_{i=1}^{n} p_i = 1 \right. \right\}, \, n \geq 2,$$

be the set of all complete finite discrete probability distributions. For all $P, Q \in \Gamma_n$, the following measures are well known in the literature on information theory and statistics:

Let us consider the measure

$$\zeta_s(P||Q) = \begin{cases} J_s(P||Q) = [s(s-1)]^{-1} \left[\sum_{i=1}^{n} \left(p_i^s q_i^{1-s} + p_i^{1-s} q_i^s \right) - 2 \right], & s \neq 0, 1 \\ J(P||Q) = \sum_{i=1}^{n} (p_i - q_i) \ln \left(\frac{p_i}{q_i} \right), & s = 0, 1 \end{cases} \tag{1}$$

for all $P, Q \in \Gamma_n$

The measure (1) can be seen in Burbea and Rao [3] and Taneja [11, 13]. These authors studied the above measure using the multiplicative constant as $(s-1)^{-1}$, while Taneja [14] studied considering it as $[s(s-1)]^{-1}$, $s \neq 0, 1$.

The expression (1) admits the following particular cases:

(i) $\zeta_{-1}(P||Q) = \zeta_2(P||Q) = \frac{1}{2}\Psi(P||Q),$
(ii) $\zeta_0(P||Q) = \zeta_1(P||Q) = J(P||Q),$
(iii) $\zeta_{1/2}(P||Q) = 8\,h(P||Q),$

where

$$\Psi(P||Q) = \chi^2(P||Q) + \chi^2(Q||P) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2(p_i + q_i)}{p_i q_i}, \tag{2}$$

and

$$h(P||Q) = \frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{p_i} - \sqrt{q_i}\right)^2, \tag{3}$$

are the *symmetric χ^2- divergence* [14, 15] and *Hellingar's discrimination* [5] respectively. More details on these divergence measures can be seen in Taneja [14].

Let us consider now the another generalized measure

$$\xi_s(P||Q) = \begin{cases} IT_s(P||Q) = [s(s-1)]^{-1}\left[\sum_{i=1}^{n}\left(\frac{p_i^s + q_i^s}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{1-s} - 1\right], & s \neq 0, 1 \\ I(P||Q) = \frac{1}{2}\left[\sum_{i=1}^{n} p_i \ln\left(\frac{2p_i}{p_i+q_i}\right) + \sum_{i=1}^{n} q_i \ln\left(\frac{2q_i}{p_i+q_i}\right)\right], & s = 1 \\ T(P||Q) = \sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\ln\left(\frac{p_i+q_i}{2\sqrt{p_i q_i}}\right), & s = 0 \end{cases} \tag{4}$$

for all $P, Q \in \Gamma_n$

The measure (4) is new in the literature and is studied for the first time by Taneja [11] and is called *arithmetic and geometric mean divergence measure*. More details on these divergence measures can be seen in Taneja [11, 13].

The measure (4) admits the following particular cases:
(i) $\xi_{-1}(P||Q) = \frac{1}{4}\Delta(P||Q).$
(ii) $\xi_1(P||Q) = I(P||Q).$
(iii) $\xi_{1/2}(P||Q) = 4\,d(P||Q).$
(iii) $\xi_0(P||Q) = T(P||Q).$
(iv) $\xi_2(P||Q) = \frac{1}{16}\Psi(P||Q).$

where

$$\Delta(P||Q) = \sum_{i=1}^{n}\frac{(p_i - q_i)^2}{p_i + q_i}, \tag{5}$$

and

$$d(P||Q) = 1 - \sum_{i=1}^{n}\left(\frac{\sqrt{p_i} + \sqrt{q_i}}{2}\right)\left(\sqrt{\frac{p_i + q_i}{2}}\right). \tag{6}$$

are the *triangular discrimination* and *d-divergence* respectively.

The symmetric divergence measures (1) and (4) admit several particular cases. An inequality among these measures [14] is given by

$$\frac{1}{4}\Delta(P||Q) \leq I(P||Q) \leq h(P||Q) \leq 4\,d(P||Q) \leq \frac{1}{8}J(P||Q) \leq T(P||Q) \leq \frac{1}{16}\Psi(P||Q). \tag{7}$$

In this paper our aim is to obtain the bounds on the probability of error in terms of the measure (1) and (4) and then establish examples of the seven measures given in (7). Some studies on divergence measures and probability of error can be seen in Taneja [10, 11].

2 Error Probability for Two Class Case

Suppose we have two pattern classes $C = \{C_1, C_2\}$ with *"a priori"* probability $p_i = \Pr\{C = C_i\}$, $i = 1, 2$. Let the feature x on X have a class-conditional probability density function $P(x|C_i)$ are known. Given a feature x on X, we can calculate the conditional (*a posteriori*) probability $P(C_i|x)$ for each i, by the Bayes rule:

$$P(C_i|x) = \Pr\{C = C_i | X = x\} = \frac{p_i P(x|C_i)}{\sum\limits_{k=1}^{2} p_k P(x|C_k)}, \; i = 1, 2.$$

It is well known that the decision rule which minimizes the probability of error is the Bayes decision rule which chooses the hypotheses (pattern classes) with the largest posterior probability. Using this rule, the partial probability of error for given $X = x$ is expressed by

$$P(e|x) = 1 - \max\{P(C_1|x), P(C_2|x)\} = \min\{P(C_1|x), P(C_2|x)\}.$$

Prior to observing X, the probability of error P_e, associated with X is defined as the expected probability of error, i.e.,

$$P_e = E_X\{P(e|x)\} = \int\limits_X P(e|x)p(x)dx,$$

where $p(x) = \sum\limits_{i=1}^{2} p_i P(x|C_i)$ is the unconditional density of X evaluated as x.

In terms of the prior probabilities for the two class case, let us consider generalized divergence measures:

$$\zeta_s = \begin{cases} J_s, & s \neq 0, 1 \\ J, & s = 0, 1 \end{cases}, \tag{8}$$

where

$$J_s = [s(s-1)]^{-1} \left[\int \left(p(x|C_1)^s p(x|C_2)^{1-s} + p(x|C_2)^s p(x|C_1)^{1-s} \right) dx - 2 \right], \; s \neq 0, 1$$

and

$$J = \int \left(P(x|C_1) - P(x|C_2) \right) \log\left(\frac{P(x|C_1)}{P(x|C_2)} \right) dx.$$

Let us consider now

$$\xi_s = \begin{cases} IT_s, & s \neq 0, 1 \\ I, & s = 1 \\ T, & s = 0 \end{cases}, \tag{9}$$

where

$$IT_s = [s(s-1)]^{-1} \left[\int \left(\frac{p(x|C_1)^s + p(x|C_2)^s}{2} \right) \left(\frac{p(x|C_1) + p(x|C_1)}{2} \right)^{1-s} dx - 1 \right], \; s \neq 0, 1$$

$$I = \int \left[\frac{P(x|C_1)}{2} \ln\left(\frac{2P(x|C_1)}{P(x|C_1) + P(x|C_2)} \right) + \frac{P(x|C_2)}{2} \ln\left(\frac{2P(x|C_2)}{P(x|C_1) + P(x|C_2)} \right) \right] dx,$$

3

and

$$T = \int \left(\frac{P(x|C_1) + P(x|C_2)}{2} \right) \ln \left(\frac{P(x|C_1) + P(x|C_2)}{2\sqrt{P(x|C_1)P(x|C_2)}} \right) dx.$$

Let us define the above measures in more general way

$$\zeta_s(p_1, p_2) = \begin{cases} J_s(p_1, p_2), & s \neq 0, 1 \\ J, (p_1, p_2) & s = 0, 1 \end{cases},$$

where

$$J_s(p_1, p_2) = [s(s-1)]^{-1} \left\{ \int \left[(p_1 p(x|C_1))^s (p_2 p(x|C_2))^{1-s} \right. \right.$$
$$\left. \left. + (p_2 p(x|C_2))^s (p_1 p(x|C_1))^{1-s} \right] dx - 1 \right\}, \ s \neq 0, 1$$

and

$$J = \int (p_1 P(x|C_1) - p_2 P(x|C_2)) \ln \left(\frac{p_1 P(x|C_1)}{p_2 P(x|C_2)} \right) dx.$$

Similarly let us define

$$\xi_s(p_1, p_2) = \begin{cases} IT_s(p_1, p_2), & s \neq 0, 1 \\ I(p_1, p_2), & s = 1 \\ T(p_1, p_2), & s = 0 \end{cases},$$

where

$$IT_s(p_1, p_2) = [s(s-1)]^{-1} \left\{ \int \left[\left(\frac{(p_1 p(x|C_1))^s + (p_2 p(x|C_2))^s}{2} \right) \cdot \right. \right.$$
$$\left. \left. \cdot \left(\frac{p_1 p(x|C_1) + p_2 p(x|C_1)}{2} \right)^{1-s} dx \right] - \frac{1}{2} \right\},$$

$$I(p_1, p_2) = \frac{1}{2} \left\{ \int \left[p_1 P(x|C_1) \ln \left(\frac{2 p_1 P(x|C_1)}{p_1 P(x|C_1) + p_1 P(x|C_2)} \right) \right. \right.$$
$$\left. \left. + p_2 P(x|C_2) \ln \left(\frac{2 p_2 P(x|C_2)}{p_1 P(x|C_1) + p_1 P(x|C_2)} \right) \right] dx \right\}$$

and

$$T(p_1, p_2) = \int \left(\frac{p_1 P(x|C_1) + p_2 P(x|C_2)}{2} \right) \ln \left(\frac{p_1 P(x|C_1) + p_2 P(x|C_2)}{2\sqrt{p_1 p_2 P(x|C_1)P(x|C_2)}} \right) dx.$$

It is easy to verify

$$\zeta_s \left(\frac{1}{2}, \frac{1}{2} \right) = \left(\frac{1}{2} \right) \zeta_s,$$

and

$$\xi_s \left(\frac{1}{2}, \frac{1}{2} \right) = \left(\frac{1}{2} \right) \xi_s,$$

4

where ζ_s and ξ_s are as given by (8) and (9) respectively.

Using Bayes rules, $p_1 P(x|C_1) = p(x)P(C_1|x)$ and $p_2 P(x|C_2) = p(x)P(C_2|x)$, we can write

$$\zeta_s(p_1, p_2) = \int p(x)\zeta_s(x)dx$$

and

$$\xi_s(p_1, p_2) = \int p(x)\xi_s(x)dx,$$

where

$$\zeta_s(x) = \begin{cases} J_s(x) = [s(s-1)]^{-1}\left[P(C_1|x)^s P(C_2|x)^{1-s} \right. \\ \left. \quad + P(C_2|x)^s P(C_2|x)^{1-s} - 1 \right], & s \neq 0, 1 \\ J(x) = (P(C_1|x) - P(C_2|x)) \ln\left(\frac{P(C_1|x)}{P(C_2|x)}\right), & s = 0, 1 \end{cases} \tag{10}$$

and

$$\xi_s(x) = \begin{cases} IT_s(x) = [s(s-1)]^{-1}\left[\left(\frac{P(C_1|x)^s + P(C_2|x)^s}{2}\right)\left(\frac{1}{2}\right)^{1-s} - \frac{1}{2}\right], & s \neq 0, 1 \\ I(x) = \frac{1}{2}\left[P(C_1|x)\ln\left(\frac{2P(C_1|x)}{P(C_1|x)+P(C_2|x)}\right) \right. \\ \left. \quad + P(C_2|x)\ln\left(\frac{2P(C_2|x)}{P(C_1|x)+P(C_2|x)}\right)\right], & s = 1 \\ T(x) = \left(\frac{P(C_1|x)+P(C_2|x)}{2}\right)\ln\left(\frac{P(C_1|x)+P(C_2|x)}{2\sqrt{P(C_1|x)P(C_2|x)}}\right), & s = 0 \end{cases}$$

2.1 Bounds on Generalized Divergence Measures

In this paper we shall obtain error bounds in terms of these generalized divergence measures. Before that we review some known results.

There exist in the literature lower bounds on P_e in terms of different divergence measures. For review refer to Taneja [10, 11]. Kailath [7] gave the following bound:

$$P_e \geq \frac{1}{4}\exp\left(-\frac{J}{2}\right),$$

where the equality holds for $J = \infty$. Later, Toussaint [17, 18] gave a tighter and general bound as

$$P_e \geq \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\exp\left[-2H(p_1, p_2) - J(p_1, p_2)\right]},$$

where

$$H(p_1, p_2) = -p_1 \log p_1 - p_2 \log p_2.$$

For $p_1 = p_2 = \frac{1}{2}$, we have

$$P_e \geq \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\exp\left(-\frac{J}{2}\right)},$$

where the equality hold for both $J = 0$ and $J = \infty$.

Toussaint [17, 18] still obtain a sharper bound on J in terms of P_e as

$$J(p_1, p_2) \geq (2P_e - 1)\log\left(\frac{P_e}{1 - P_e}\right).$$

When $p_1 = p_2 = \frac{1}{2}$, one gets

$$J \geq 2(2P_e - 1)\log\left(\frac{P_e}{1 - P_e}\right).$$

Theorem 2.1. *We have*

$$\zeta_s(p_1, p_2) \geq \zeta_s(P_e),$$

where

$$\zeta_s(P_e) = \begin{cases} J_s(P_e) = [s(s-1)]^{-1}\left[P_e^s(1-P_e)^{1-s} + (1-P_e)^s P_e^{1-s} - 1\right] & s \neq 0, 1 \\ J(P_e) = (2P_e - 1)\ln\left(\frac{P_e}{1-P_e}\right), & s = 0, 1 \end{cases}.$$

Proof. It is sufficient to show the result only for $s \neq 0, 1$, since for $s = 0, 1$ the result is already known in the literature or follows in view of continuity with respect parameter s.

We have

$$P_e(x) = \min\left\{P(C_1|x),\, P(C_2|x)\right\}.$$

As $\zeta_s(x)$ given by (10) is symmetric with respect to $P(C_1|x)$ and $P(C_2|x)$, let us consider $P(C_1|x) = P_e(x)$ and $P(C_2|x) = 1 - P_e(x)$. Making these substitutions in (10), we get

$$\zeta_s(x) = \begin{cases} J_s(x) = [s(s-1)]^{-1}\left[P_e(x)^s\,(1 - P_e(x))^{1-s}\,(1 - P_e(x))^s\,P_e(x)^{1-s} - 1\right], & s \neq 0, 1 \\ J(x) = (2P_e(x) - 1)\ln\left(\frac{P_e(x)}{1-P_e(x)}\right), & s = 0, 1 \end{cases}. \tag{11}$$

Now consider a function

$$f_s(a) = \begin{cases} [s(s-1)]^{-1}\left[a^s\,(1-a)^{1-s} + (1-a))^s\,a^{1-s} - 1\right], & s \neq 0, 1 \\ (2a - 1)\ln\left(\frac{a}{1-a}\right), & s = 0, 1 \end{cases}$$

or, simply,

$$f_s(a) = [s(s-1)]^{-1}\left[a^s\,(1-a)^{1-s} + (1-a))^s\,a^{1-s} - 1\right], \quad s \neq 0, 1.$$

We have

$$f_s'(a) = [s(s-1)a(1-a)]^{-1}\left[(s-1)a^s\,(1-a)^{1-s}\right.$$
$$\left. + (1-s)a^{1-s}(1-a) - (1-a)^s a^{2-s}\right]$$

and

$$f_s''(a) = a^{s-2}(1-a)^{-s-1} + (1-a)^{s-2}a^{-s-1}$$

for all $0 < a < \frac{1}{2}$.

This gives $f_s''(a) \geq 0$, for any $s \neq 0, 1$. For $s = 0, 1$ the result follows by continuity, i.e, $f_s(a)$ is convex function of a $(0 < a < \frac{1}{2})$ for any $s \in (-\infty, \infty)$.

Taking expectation on both sides of (11) and using Jensen's inequality we get the required result. $\qquad\square$

Corollary 2.1. *When $p_1 = p_2 = \frac{1}{2}$, we have*

$$\zeta_s \geq \begin{cases} 2[s(s-1)]^{-1} \left[P_e^s (1-P_e)^{1-s} + (1-P_e)^s P_e^{1-s} - 1 \right] & s \neq 0, 1 \\ 2(2P_e - 1) \ln\left(\frac{P_e}{1-P_e} \right), & s = 0, 1 \end{cases}$$

Theorem 2.2. *We have*

$$\xi_s(p_1, p_2) \geq \xi_s(P_e),$$

where

$$\xi_s(P_e) = \begin{cases} IT_s(P_e) = [s(s-1)]^{-1} \left[\left(\frac{P_e^s + (1-P_e)^s}{2} \right) \left(\frac{1}{2} \right)^{1-s} - \frac{1}{2} \right], & s \neq 0, 1 \\ I(P_e) = \frac{1}{2} \left[P_e \ln\left(\frac{2P_e}{2P_e - 1} \right) + (1-P_e) \ln\left(\frac{2(1-P_e)}{2P_e - 1} \right) \right], & s = 1 \\ T(P_e) = \left(\frac{2P_e - 1}{2} \right) \ln\left(\frac{2P_e - 1}{2\sqrt{P_e(1-P_e)}} \right), & s = 0 \end{cases} .$$

Proof. Following the lines of Theorem 2.1, we can write

$$\xi_s(x) = \begin{cases} IT_s(x) = [s(s-1)]^{-1} \left[\left(\frac{P_e(x)^s + (1-P_e(x))^s}{2} \right) \left(\frac{1}{2} \right)^{1-s} - \frac{1}{2} \right], & s \neq 0, 1 \\ I(x) = \frac{1}{2} \left[P_e(x) \ln\left(\frac{2P_e(x)}{2P_e(x) - 1} \right) + (1-P_e(x)) \ln\left(\frac{2(1-P_e(x))}{2P_e(x) - 1} \right) \right] & s = 1 \\ T(x) = \left(\frac{2P_e(x) - 1}{2} \right) \ln\left(\frac{2P_e(x) - 1}{2\sqrt{P_e(x)(1-P_e(x))}} \right), & s = 0 \end{cases} . \qquad (12)$$

Now consider a function

$$g_s(a) = \begin{cases} IT_s(a) = [s(s-1)]^{-1} \left[\left(\frac{a^s + (1-a)^s}{2} \right) \left(\frac{1}{2} \right)^{1-s} - \frac{1}{2} \right], & s \neq 0, 1 \\ I(a) = \frac{1}{2} \left[a \ln\left(\frac{2a}{2a-1} \right) + (1-a) \ln\left(\frac{2(1-a)}{2a-1} \right) \right], & s = 1 \\ T(a) = \left(\frac{2a-1}{2} \right) \ln\left(\frac{2a-1}{2\sqrt{a(1-a)}} \right), & s = 0 \end{cases} ,$$

for all $0 < a < \frac{1}{2}$.

We have

$$g'_s(a) = \begin{cases} IT'_s(a) = \frac{2^{s-2}[a^s(a-1) + a(1-a)^s]}{(s-1)(a-1)}, & s \neq 0, 1 \\ I'(a) = \ln\left(\frac{a}{1-a} \right), & s = 1 \\ T'(a) = \frac{2a-1}{4a(1-a)}, & s = 0 \end{cases}$$

and

$$g''_s(a) = \begin{cases} IT''_s(a) = 2^{s-2} \left[\frac{a^s(1-a)^2 + a^2(1-a)^s}{a^2(1-a)^2} \right], & s \neq 0, 1 \\ I''(a) = \frac{1}{2a(1-a)}, & s = 1 \\ T''(a) = \frac{a^2 + (1-a)^2}{4a^2(1-a)^2}, & s = 0 \end{cases} .$$

This gives $g''_s(a) \geq 0$, i.e, $g_s(a)$ is convex function of a $(0 < a < \frac{1}{2})$ for any $s \in (-\infty, \infty)$.

Taking expectation on both sides of (12) and using Jensen's inequality we get the required result. □

Corollary 2.2. *When $p_1 = p_2 = \frac{1}{2}$, we have*

$$\xi_s \geq \begin{cases} 2\left[s(s-1)\right]^{-1}\left[\left(\frac{P_e^s+(1-P_e)^s}{2}\right)\left(\frac{1}{2}\right)^{1-s} - \frac{1}{2}\right], & s \neq 0,1 \\ P_e \ln\left(\frac{2P_e}{2P_e-1}\right) + (1-P_e)\ln\left(\frac{2(1-P_e)}{2P_e-1}\right), & s = 1 \\ (2P_e-1)\ln\left(\frac{2P_e-1}{2\sqrt{P_e(1-P_e)}}\right), & s = 0 \end{cases}$$

3 $f-$Divergence and Probability of Error

Csiszár [4] have given a measure for the divergence between two probability density functions, say $p(x)$ and $q(x)$. This so called $f-$divergence given by

$$C_f(p,q) = \int_X f\left(\frac{p(x)}{q(x)}\right)q(x)dx. \tag{13}$$

The function $f(x)$, with $x \in (0,\infty)$ is a convex function which has to satisfy the conditions

$$f(0) = \lim_{u\downarrow 0} f(u);\; 0f\left(\frac{0}{0}\right) = 0\,;\; 0f\left(\frac{a}{0}\right) = \lim_{\in\downarrow\infty}\in f\left(\frac{a}{\in}\right) = a\lim_{x\to\infty}\frac{f(u)}{u}. \tag{14}$$

It can be easily checked that $C_f(p,q) \geq f(1)$ and that $C_f(p,q) = f(1)$ only when $p(x) = q(x)$ a.e. Thus, $C_f(p,q) - f(1)$ is a distance or divergence measure in the sense that $C_f(p,q) - f(1) \geq 0$. However, it is not symmetric in p and q and in general does not satisfy triangle inequality. Some applications of $f-$divergence in connection with divergence measures are given in Taneja and Kumar [12].

Boekee and Van der Lubbe [2] have introduced the average $f-$divergence between two hypothesis C_1 and C_2 in terms of their "a posteriori" probabilities. This average $f-$divergence is defined as

$$\overline{C}_f(C_1, C_2) = \int_X f\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)p(x)dx = E_X\left\{f\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}. \tag{15}$$

If we introduce the function $f_*(u) = u\,f\left(\frac{1-u}{u}\right)$ and set $u = u(x) = P(C_2|x)$, it is easy to see from $P(C_1|x) = 1 - P(C_2|x)$ that

$$\overline{C}_f(C_1, C_2) = E_X\left\{f_*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}$$

From Vajda [19] it follows that $f_*(u)$ is convex on $[0,1]$ and is strictly convex iff $f(u)$ is strictly convex.

3.1 A Class of Upper Bounds

In [2] it has been shown that the Bayesian probability of error can be upper bounds in terms of the average $f-$divergence $\overline{C}_f(C_1, C_2)$. This upper bound is given by

$$P_e \leq \frac{f_0 P(C_2) + f_\infty P(C_1) - \overline{C}_f(C_1, C_2)}{f_2 - f_1}, \tag{16}$$

8

where f_2 should be finite with $f_0 = \lim\limits_{u \downarrow \infty} f(u)$; $f_1 = f(1)$; $f_2 = f_0 + f_\infty$; $f_\infty = \lim\limits_{u \to \infty} \frac{f(u)}{u}$.

The above bound is valid only for *every convex function* $f(u)$ which satisfies the conditions given in (14). However, if $f_*(u) = u f \left(\frac{1-u}{u} \right)$ is symmetric with respect to $u = \frac{1}{2}$ i.e., $f_*(u) = f_*(1-u)$, $\forall u \in (0,1)$, the bound (16) can be written in a simpler form given by

$$P_e \le \frac{1}{2f_\infty - f_1} \left[f_\infty - \overline{C}_f(C_1, C_2) \right], \tag{17}$$

provided f_∞ is finite and $f_*(u) = f_*(1-u)$, $\forall u \in (0,1)$.

As a consequence of result (17) we have the following two theorems in terms of measures given in (1) and (4) respectively.

Theorem 3.1. *The following bound holds*

$$P_e \le \frac{1}{2} \left[1 + s(s-1)\overline{\zeta}_s(C_1, C_2) \right], \, 0 < s < 1, \tag{18}$$

where

$$\overline{\zeta}_s(C_1, C_2) = E_X \left\{ f_{\zeta_s}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\}$$

with

$$f_{\zeta_s}^*(x) = \begin{cases} [s(s-1)]^{-1} \left[(1-x)^{1-s} x^s + x^{1-s}(1-x)^s - 1 \right], & s \ne 0,1 \\ (2x-1) \ln \left(\frac{x}{1-x} \right) & s = 0,1 \end{cases}$$

for all $x \in (0,1)$.

Proof. Let us consider

$$\overline{\zeta}_s(C_1, C_2) = E_X \left\{ f_{\zeta_s}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$f_{\zeta_s}^*(x) = \begin{cases} x f_{J_s} \left(\frac{1-x}{x} \right) = \frac{(1-x)^{1-s} x^s + x^{1-s}(1-x)^s - 1}{s(s-1)}, & s \ne 0,1 \\ x f_J \left(\frac{1-x}{x} \right) = (2x-1) \ln \left(\frac{x}{1-x} \right), & s = 0,1 \end{cases} \tag{19}$$

with

$$f_{\zeta_s}(x) = \begin{cases} f_{J_s}(x) = [s(s-1)]^{-1} \left[x^s + x^{1-s} - (x+1) \right], & s \ne 0,1 \\ f_J(x) = (x-1) \ln x & s = 0,1 \end{cases} \tag{20}$$

for all $x \in (0,1)$.

The convexity of the function $f_{\zeta_s}(x)$ given (20) can be seen in Taneja [14]. From the expression (19), we observe that

$$f_{\zeta_s}^*(x) = f_{\zeta_s}^*(1-x). \tag{21}$$

Also, we have

$$f_{\zeta_\infty} = \lim_{x \to \infty} \frac{f_\zeta(x)}{x} = \begin{cases} \frac{-1}{s(s-1)}, & 0 < s < 1 \\ \infty & \text{otherwise} \end{cases} \tag{22}$$

and

$$f_\zeta(1) = 0. \tag{23}$$

Expression (17) together with (21)-(23) give the required result (18). $\qquad \square$

Theorem 3.2. *The following bound holds*

$$P_e \leq \frac{1}{2}\left[1 - \frac{2s(s-1)}{2^{-s}-1}\overline{\xi}_s(C_1, C_2)\right], \ s < 1 \tag{24}$$

where

$$\overline{\xi}_s(C_1, C_2) = E_X\left\{f_{\xi_s}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}$$

with

$$f_{\xi_s}^*(x) = \begin{cases} f_{IT_s}^*(x) = [2s(s-1)]^{-1}\left[2^{-s}\left((1-x)^{1-s}+x^{1-s}\right)-1\right], & s \neq 0, 1 \\ f_I^*(x) = \frac{1}{2}\left[\ln 2 + x\ln x + (1-x)\ln(1-x)\right], & s = 0 \\ f_T^*(x) = \frac{1}{2}\ln\left(\frac{1}{2\sqrt{x(1-x)}}\right), & s = 1 \end{cases}$$

for all $x \in (0, 1)$.

Proof. Let us write

$$\overline{\xi}_s(C_1, C_2) = E_X\left\{f_{\xi_s}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

where

$$f_{\xi_s}^*(x) = \begin{cases} x\, f_{IT_s}\left(\frac{1-x}{x}\right) = \frac{2^{-s}\left((1-x)^{1-s}+x^{1-s}\right)-1}{2s(s-1)}, & s \neq 0, 1 \\ x\, f_I\left(\frac{1-x}{x}\right) = \frac{1}{2}\left[\ln 2 + x\ln x + (1-x)\ln(1-x)\right], & s = 0 \\ x\, f_T\left(\frac{1-x}{x}\right) = \frac{1}{2}\ln\left(\frac{1}{2\sqrt{x(1-x)}}\right), & s = 1 \end{cases} \tag{25}$$

with

$$f_{\xi_s}(x) = \begin{cases} f_{IT_s}(x) = [s(s-1)]^{-1}\left[\left(\frac{x^{1-s}+1}{2}\right)\left(\frac{x+1}{2}\right)^s - \left(\frac{x+1}{2}\right)\right], & s \neq 0, 1 \\ f_I(x) = \frac{x}{2}\ln x + \left(\frac{x+1}{2}\right)\ln\left(\frac{x+1}{2}\right), & s = 0 \\ f_T(x) = \frac{x+1}{2}\ln x\left(\frac{x+1}{2\sqrt{x}}\right), & s = 1 \end{cases} \tag{26}$$

for all $x \in (0, 1)$.

The convexity of the function $f_{\xi_s}(x)$ given (26) can be seen in Taneja [13]. From the expression (25) we observe that

$$f_{\xi_s}^*(x) = f_{\xi_s}^*(1 - x). \tag{27}$$

Let us calculate now f_{ξ_∞}. We have

$$\lim_{x\to\infty}\frac{f_{IT_s}(x)}{x} = \lim_{x\to\infty}[s(s-1)]^{-1}\frac{1}{x}\left[\left(\frac{x^{1-s}+1}{2}\right)\left(\frac{x+1}{2}\right)^s - \left(\frac{x+1}{2}\right)\right]$$

$$= \lim_{x\to\infty}[s(s-1)]^{-1}\left[\left(\frac{x^{1-s}+1}{2x}\right)\left(\frac{x+1}{2}\right)^s - \left(\frac{1}{2}+\frac{1}{x}\right)\right]$$

$$= \lim_{x\to\infty}[s(s-1)]^{-1}\left[2^{-1-s}\left(\frac{x^{1-s}+1}{x}\right)(x+1)^s - \left(\frac{1}{2}+\frac{1}{x}\right)\right]$$

$$= \lim_{x\to\infty}[s(s-1)]^{-1}\left[2^{-1-s}\left(\left(\frac{x+1}{x}\right)^s + \frac{(x+1)^s}{x}\right) - \left(\frac{1}{2}+\frac{1}{x}\right)\right]$$

$$= [s(s-1)]^{-1}\left[2^{-1-s}\left(\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^s + \lim_{x\to\infty}\frac{(x+1)^s}{x}\right) - \lim_{x\to\infty}\left(\frac{1}{2}+\frac{1}{x}\right)\right]$$

$$= [2s(s-1)]^{-1}\left[2^{-s}\left(1 + \lim_{x\to\infty}\frac{(x+1)^s}{x}\right) - 1\right], \ s \neq 0, 1. \tag{28}$$

10

We can easily verify that

$$\lim_{x \to \infty} \frac{(x+1)^s}{x} \begin{cases} 0, & s < 1 \\ 1, & s = 1 \ . \\ \infty, & s > 1 \end{cases} \tag{29}$$

Expressions (28) and (29) together give

$$f_{\xi_\infty} = \lim_{x \to \infty} \frac{f_\xi(x)}{x} = \begin{cases} [2s(s-1)]^{-1} [2^{-s} - 1], & s < 1, \ s \neq 0 \\ \frac{1}{2} \ln 2, & s = 0 \\ \infty, & s \geq 1 \end{cases} \tag{30}$$

Also we have

$$f_\xi(1) = 0. \tag{31}$$

Expression (17) together with (27), (30) and (31) give the required result (24). □

Here below are the particular cases of the Theorems 3.1 and 3.2. These can also be obtained directly using result (17). These examples are the seven measures appearing (7).

Example 3.1. *Let us consider the measure*

$$\overline{J}(C_1, C_2) = E_X \left\{ f_J^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$f_J^*(x) = (2x - 1) \ln \left(\frac{x}{1-x} \right), \ \forall x \in (0, 1).$$

We can't obtain the upper bound on the probability of error in terms of J-divergence, since according to (22), f_{J_∞} is infinite.

Example 3.2. *In particular for $s = \frac{1}{2}$ in (18), we have the following upper bound on the probability of error in terms of the Hellinger distance:*

$$P_e \leq \frac{1}{2} \left[1 - 2\overline{h}(C_1, C_2) \right],$$

where

$$\overline{h}(C_1, C_2) = E_X \left\{ f_h^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

with

$$f_h^*(x) = \frac{1 - 2\sqrt{x(1-x)}}{2}, \ \forall x \in (0, 1).$$

Example 3.3. *In particular for $s = 0$ in (24), we have the following upper bound on the probability of error in terms of the Jensen-Shannon divergence:*

$$P_e \leq \frac{1}{2} \left[1 - \frac{2}{\ln 2} \overline{I}(C_1, C_2) \right],$$

where

$$\overline{I}(C_1, C_2) = E_X \left\{ f_I^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

11

with

$$f_I^*(x) = \frac{1}{2}\left[\ln 2 + x\ln x + (1-x)\ln(1-x)\right], \; \forall x \in (0,1).$$

Example 3.4. *Let us consider the measure*

$$\overline{T}(C_1, C_2) = E_X\left\{f_T^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

where

$$f_T^*(x) = \frac{1}{2}\ln\left(\frac{1}{2\sqrt{x(1-x)}}\right), \; \forall x \in (0,1).$$

We can't obtain the upper bound on the probability of error in terms of the arithmetic and geometric mean divergence, since according to (30), f_{T_∞} is infinite in this case.

Example 3.5. *In particular for $s = \frac{1}{2}$ in (24), we have the following upper bound on the probability of error in terms of d-divergence:*

$$P_e \leq \frac{1}{2}\left[1 - \left(\frac{4}{2-\sqrt{2}}\right)\overline{d}(C_1, C_2)\right],$$

where

$$\overline{d}(C_1, C_2) = E_X\left\{f_d^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

with

$$f_d^*(x) = \frac{1}{2} - \sqrt{2}\left(\sqrt{x} + \sqrt{1-x}\right), \; \forall x \in (0,1).$$

Example 3.6. *In particular for $s = -1$ in (24), we have the following upper bound on the probability of error in terms of triangular discrimination:*

$$P_e \leq \frac{1}{2}\left[1 - \overline{\Delta}(C_1, C_2)\right],$$

where

$$\overline{\Delta}(C_1, C_2) = E_X\left\{f_\Delta^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}$$

with

$$f_\Delta^*(x) = (2x-1)^2, \; \forall x \in (0,1).$$

Example 3.7. *Let us consider the measure*

$$\overline{\Psi}(C_1, C_2) = E_X\left\{f_\Psi^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

where

$$f_\Psi^*(x) = \frac{(2x-1)^2}{x(1-x)}, \; \forall x \in (0,1).$$

We can't obtain the upper bound on the probability of error in terms of symmetric χ^2-divergence, since according to (23), f_{Ψ_∞} is infinite for $s = 2$.

4   Difference of Divergence Measures

From the Examples 3.1-3.7 we observe that there are only four measures that give upper bounds on the probability of error. According to (1), these four measures bear the following inequality:

$$\frac{1}{4}\Delta(P||Q) \leq I(P||Q) \leq h(P||Q) \leq 4\,d(P||Q). \tag{32}$$

The inequality (32) admits the following nonnegative differences

$$D_{d\Delta}(P||Q) = 4\,d(P||Q) - \frac{1}{4}\Delta(P||Q), \tag{33}$$

$$D_{dh}(P||Q) = 4\,d(P||Q) - h(P||Q), \tag{34}$$

$$D_{dI}(P||Q) = 4\,d(P||Q) - I(P||Q), \tag{35}$$

$$D_{hI}(P||Q) = h(P||Q) - I(P||Q), \tag{36}$$

$$D_{h\Delta}(P||Q) = h(P||Q) - \frac{1}{4}\Delta(P||Q) \tag{37}$$

and

$$D_{I\Delta}(P||Q) = I(P||Q) - \frac{1}{4}\Delta(P||Q). \tag{38}$$

The above six measures can be related by the following (ref. Taneja [16]):

$$D_{I\Delta}(P||Q) \leq \frac{2}{3}D_{h\Delta}(P||Q) \leq \frac{8}{15}D_{d\Delta}(P||Q) \leq$$
$$\leq \frac{8}{3}D_{dh}(P||Q) \leq \frac{8}{7}D_{dI}(P||Q) \leq 2\,D_{hI}(P||Q). \tag{39}$$

Here below we shall give bounds on the probability of error in terms of the measures (33)-(38).

Result 4.1. *The following bound holds*

$$P_e \leq \frac{1}{2}\left[1 - \frac{4}{7 - 4\sqrt{2}}\overline{D}_{d\Delta}(C_1, C_2)\right], \tag{40}$$

where

$$\overline{D}_{d\Delta}(C_1, C_2) = E_X\left\{f_{d\Delta}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}$$

with

$$f_{d\Delta}^*(x) = \frac{7}{4} - \sqrt{2}\left(\sqrt{x} + \sqrt{1-x}\right) + x(1-x), \forall x \in (0,1).$$

Proof. Let us consider

$$\overline{D}_{d\Delta}(C_1, C_2) = E_X\left\{f_{d\Delta}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

where

$$f_{d\Delta}^*(x) = x\,f_{d\Delta}\left(\frac{1-x}{x}\right) = \frac{7}{4} - \sqrt{2}\left(\sqrt{x} + \sqrt{1-x}\right) + x(1-x),$$

with

$$f_{d\Delta}(x) = 2x + 2 - \left(\sqrt{x} + 1\right)\sqrt{2x + 2} - \frac{(x-1)^2}{4(x+1)^2}.$$

Now we shall prove the convexity of the function $f_{d\Delta}(x)$. We have

$$f'_{d\Delta}(x) = 2 - \frac{\sqrt{2x+2}}{2\sqrt{x}} - \frac{1+\sqrt{x}}{\sqrt{2x+2}} - \frac{x-1}{2(x+1)^2} + \frac{(x-1)^2}{4(x+1)^2}$$

and

$$f''_{d\Delta}(x) = \frac{(x^{3/2}+1)(x+1)^2 - 4x^{3/2}\sqrt{2x+2}}{2x^{3/2}(x+1)^3\sqrt{2x+2}}$$

$$= \frac{8}{2x^{3/2}(x+1)^3\sqrt{2x+2}}\sqrt{\frac{x+1}{2}}\left[\left(\frac{x^{3/2}+1}{2}\right)\left(\frac{x+1}{2}\right)^{3/2} - x^{3/2}\right], \forall x > 0. \quad (41)$$

Now we shall prove the non-negativity of the expression (41). We can easily check that

$$\left(\frac{\sqrt{x}+1}{2}\right)^3 \leq \frac{x^{3/2}+1}{2}. \quad (42)$$

On the other side we know that [13] pp. 279:

$$x^{3/2} \leq \left(\frac{\sqrt{x}+1}{2}\right)^3\left(\frac{x+1}{2}\right)^{3/2}. \quad (43)$$

Expressions (42) and (43) together give

$$x^{3/2} \leq \left(\frac{\sqrt{x}+1}{2}\right)^3\left(\frac{x+1}{2}\right)^{3/2} \leq \left(\frac{x^{3/2}+1}{2}\right)\left(\frac{x+1}{2}\right)^{3/2}. \quad (44)$$

Expression (44) proves the non-negativity of the expression (41), i.e., $f''_{d\Delta}(x) \geq 0$, $\forall x > 0$. This proves the convexity of the function $f_{d\Delta}(x)$, $\forall x > 0$.

We have

$$f^*_{d\Delta}(x) = f^*_{d\Delta}(1 - x), \ x \in (0, 1), \quad (45)$$

$$f_{(d\Delta)_\infty} = \lim_{x\to\infty}\frac{f_{d\Delta}(x)}{x} = \frac{7}{4} - \sqrt{2} = \frac{7 - 4\sqrt{2}}{4} \quad (46)$$

and

$$f_1 = f_{d\Delta}(1) = 0. \quad (47)$$

The result (17) together with the expressions (45)-(47) gives the required result (40). $\square$

Result 4.2. *The following bound holds*

$$P_e \leq \frac{1}{2}\left[1 - \frac{2}{4 - 2\sqrt{2} - \ln 2}\overline{D}_{dI}(C_1, C_2)\right], \quad (48)$$

14

where

$$\overline{D}_{dI}(C_1, C_2) = E_X \left\{ f_{dI}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\}$$

with

$$f_{dI}^*(x) = 2 - 2\sqrt{2} \left(\sqrt{x} + \sqrt{1-x} \right) - \frac{1}{2} \left[\ln 2 + x \ln x + (1-x) \ln(1-x) \right], \, \forall x \in (0, 1).$$

Proof. Let us consider

$$\overline{D}_{dI}(C_1, C_2) = E_X \left\{ f_{dI}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$f_{dI}^*(x) = x \, f_{dI} \left(\frac{1-x}{x} \right)$$

$$= 2 - 2\sqrt{2} \left(\sqrt{x} + \sqrt{1-x} \right) - \frac{1}{2} \left[\ln 2 + x \ln x + (1-x) \ln(1-x) \right]$$

with

$$f_{dI}(x) = 2x + 2 - \left(\sqrt{x} + 1 \right) \sqrt{2x+2} - \frac{1}{2} x \ln x - \frac{x+1}{2} \ln \left(\frac{2}{x+1} \right).$$

Now we shall prove the convexity of the function $f_{dI}(x)$. We have

$$f_{dI}'(x) = 2 - \frac{\sqrt{2x+2}}{2\sqrt{x}} - \frac{1+\sqrt{x}}{\sqrt{2x+2}} - \frac{1}{2} \ln x - \frac{1}{2} \ln \left(\frac{2}{x+1} \right)$$

and

$$f_{dI}''(x) = \frac{x^{3/2} + 1 - \sqrt{x}\sqrt{2x+2}}{2x^{3/2}(x+1)\sqrt{2x+2}}$$

$$= \frac{1}{x^{3/2}(x+1)\sqrt{2x+2}} \left[\left(\frac{x^{3/2}+1}{2} \right) - \sqrt{x}\sqrt{\frac{x+1}{2}} \right], \, \forall x > 0. \qquad (49)$$

Now we shall prove the non-negativity of the expression (49). We know that [14], pp. 358:

$$\sqrt{x} \le \left(\frac{\sqrt{x}+1}{2} \right)^2 \text{ and } \left(\frac{\sqrt{x}+1}{2} \right) \sqrt{\frac{x+1}{2}} \le \left(\frac{x+1}{2} \right).$$

This give

$$\sqrt{x}\sqrt{\frac{x+1}{2}} \le \left(\frac{\sqrt{x}+1}{2} \right)^2 \sqrt{\frac{x+1}{2}} \le \left(\frac{x+1}{2} \right) \left(\frac{\sqrt{x}+1}{2} \right). \qquad (50)$$

By simple calculations, we can check that

$$\left(\frac{x+1}{2} \right) \left(\frac{\sqrt{x}+1}{2} \right) \le \frac{x^{3/2}+1}{2}. \qquad (51)$$

The expressions (50) and (51) together give the non-negativity of the expression (49), i.e., $f_{dI}''(x) \ge 0$, $\forall x > 0$. This proves the convexity of the function $f_{dI}(x)$, $\forall x > 0$.

We have

$$f_{d\Delta}^*(x) = f_{d\Delta}^*(1-x), \ x \in (0,1), \tag{52}$$

$$f_{(d\Delta)_\infty} = \lim_{x \to \infty} \frac{f_{d\Delta}(x)}{x} = 2 - \sqrt{2} - \frac{1}{2} \ln 2 = \frac{2}{4 - 2\sqrt{2} - \ln 2} \tag{53}$$

and

$$f_1 = f_{d\Delta}(1) = 0. \tag{54}$$

The result (17) together with the expressions (52)-(54) gives the required result (48). $\quad\square$

Result 4.3. *The following bound holds*

$$P_e \leq \frac{1}{2} \left[1 - \frac{2}{3 - 2\sqrt{2}} \overline{D}_{dh}(C_1, C_2) \right], \tag{55}$$

where

$$\overline{D}_{dh}(C_1, C_2) = E_X \left\{ f_{dh}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\}$$

with

$$f_{dh}^*(x) = \frac{3}{2} - \sqrt{2} \left(\sqrt{x} + \sqrt{1-x} \right) + \frac{1}{2} \sqrt{x(1-x)}, \ \forall x \in (0,1).$$

Proof. Let us consider

$$\overline{D}_{dh}(C_1, C_2) = E_X \left\{ f_{dh}^* \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$f_{dh}^*(x) = x f_{dh} \left(\frac{1-x}{x} \right) = \frac{3}{2} - \sqrt{2} \left(\sqrt{x} + \sqrt{1-x} \right) + \frac{1}{2} \sqrt{x(1-x)}$$

with

$$f_{dh}(x) = 2x + 2 - \left(\sqrt{x} + 1 \right) \sqrt{2x+2} - \frac{\left(\sqrt{x} - 1 \right)^2}{2}.$$

Now we shall prove the convexity of the function $f_{dh}(x)$. We have

$$f_{dh}'(x) = 2 - \frac{\sqrt{2x+2}}{2\sqrt{x}} - \frac{1 + \sqrt{x}}{\sqrt{2x+2}} - \frac{\sqrt{x} - 1}{2\sqrt{x}}$$

and

$$f_{dh}''(x) = \frac{2(x^{3/2} + 1) - (x+1)\sqrt{2x+2}}{4x^{3/2}(x+1)\sqrt{2x+2}}$$

$$= \frac{1}{x^{3/2}(x+1)\sqrt{2x+2}} \left[\frac{x^{3/2} + 1}{2} - \left(\frac{x+1}{2} \right)^{3/2} \right], \ \forall x > 0, \tag{56}$$

The non-negativity of the expression (56) follows from the fact that the function $\left(\frac{x^s+1}{2} \right)^{1/s}$, $s \neq 0$ is monotonically increasing function of s. For proof refer to Beckenbach and Bellman [1].

16

Thus, we have $f''_{dh}(x) \geq 0$, $\forall x > 0$, consequently proving the convexity of the function $f_{dh}(x)$, $\forall x > 0$.

We have

$$f^*_{dh}(x) = f^*_{dh}(1 - x), \ x \in (0, 1), \tag{57}$$

$$f_{(dh)_\infty} = \lim_{x \to \infty} \frac{f_{dh}(x)}{x} = \frac{3}{2} - \sqrt{2} = \frac{3 - 2\sqrt{2}}{2} \tag{58}$$

and

$$f_1 = f_{dh}(1) = 0. \tag{59}$$

The result (17) together with the expressions (57)-(59) gives the required result (55). $\qquad\square$

Result 4.4. *The following bound holds*

$$P_e \leq \frac{1}{2} \left[1 - 4 \overline{D}_{h\Delta}(C_1, C_2) \right], \tag{60}$$

where

$$\overline{D}_{h\Delta}(C_1, C_2) = E_X \left\{ f^*_{h\Delta} \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\}$$

with

$$f^*_{h\Delta}(x) = \frac{1}{4} + x(1 - x) - \sqrt{x(1 - x)}, \ \forall x \in (0, 1).$$

Proof. Let us consider

$$\overline{D}_{h\Delta}(C_1, C_2) = E_X \left\{ f^*_{h\Delta} \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$f^*_{h\Delta}(x) = x \, f_{h\Delta} \left(\frac{1 - x}{x} \right) = \frac{1}{4} + x(1 - x) - \sqrt{x(1 - x)}$$

with

$$f_{h\Delta}(x) = \frac{(\sqrt{x} - 1)^2}{2} - \frac{(x - 1)^2}{4(x - 1)}.$$

Now we shall prove the convexity of the function $f_{h\Delta}(x)$. We have

$$f'_{h\Delta}(x) = \frac{\sqrt{x} - 1}{2\sqrt{x}} - \frac{x - 1}{2(x + 1)} + \frac{(x - 1)^2}{4(x + 1)^2}$$

and

$$\begin{aligned}
f''_{h\Delta}(x) &= \frac{x^4 + 3x^2 + 3x + x - 8x^{5/2}}{4x^{5/2}(x + 1)^3} \\
&= \frac{(x + 1)^3 - x^{3/2}}{4x^{3/2}(x + 1)^3} \\
&= \frac{(\sqrt{x} - 1)^2 \left[(\sqrt{x} + 1)^2 (x + 1) + 4x \right]}{4x^{3/2}(x + 1)^3} \geq 0, \ \forall x > 0. \tag{61}
\end{aligned}$$

17

Since, $f''_{h\Delta}(x) \geq 0$, $\forall x > 0$ proving the convexity of the function $f_{h\Delta}(x)$, $\forall x > 0$.
We have

$$f^*_{h\Delta}(x) = f^*_{h\Delta}(1-x), \; x \in (0,1), \tag{62}$$

$$f_{(h\Delta)_\infty} = \lim_{x \to \infty} \frac{f_{h\Delta}(x)}{x} = \frac{1}{4} \tag{63}$$

and

$$f_1 = f_{h\Delta}(1) = 0. \tag{64}$$

The result (17) together with the expressions (62)-(64) gives the required result (60). $\quad\square$

Result 4.5. *The following bound holds*

$$P_e \leq \frac{1}{2} \left[1 - \frac{2}{1 - \ln 2} \overline{D}_{hI}(C_1, C_2) \right], \tag{65}$$

where

$$\overline{D}_{hI}(C_1, C_2) = E_X \left\{ f^*_{hI} \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\}$$

with

$$f^*_{hI}(x) = \frac{1}{2} \left[1 - 2\sqrt{x(1-x)} - (\ln 2 + x \ln x + (1-x)\ln(1-x)) \right], \; \forall x \in (0,1).$$

Proof. Let us consider

$$\overline{D}_{hI}(C_1, C_2) = E_X \left\{ f^*_{hI} \left(\frac{P(C_1|x)}{P(C_2|x)} \right) P(C_1|x) \right\},$$

where

$$
\begin{aligned}
f^*_{hI}(x) &= x \, f_{hI} \left(\frac{1-x}{x} \right) \\
&= \frac{1}{2} \left[1 - 2\sqrt{x(1-x)} - (\ln 2 + x \ln x + (1-x)\ln(1-x)) \right]
\end{aligned}
$$

with

$$f_{hI}(x) = \frac{(\sqrt{x} - 1)^2}{2} - \frac{x}{2} \ln x - \frac{x+1}{2} \ln \left(\frac{2}{x+1} \right).$$

Now we shall prove the convexity of the function $f_{hI}(x)$. We have

$$f'_{hI}(x) = \frac{\sqrt{x} - 1}{2\sqrt{x}} - \frac{1}{2} \ln x - \frac{1}{2} \ln \left(\frac{2}{x+1} \right)$$

and

$$f''_{hI}(x) = \frac{(\sqrt{x} - 1)^2}{4x^{3/2}(x+1)} \geq 0, \; \forall x > 0. \tag{66}$$

Since, $f''_{hI}(x) \geq 0$, $\forall x > 0$ proving the convexity of the function $f_{hI}(x)$, $\forall x > 0$.

18

We have

$$f_{hI}^*(x) = f_{h\Delta}^*(1-x),\ x \in (0,1),\tag{67}$$

$$f_{(hI)_\infty} = \lim_{x\to\infty} \frac{f_{hI}(x)}{x} = \frac{1}{2} - \frac{1}{2}\ln 2 = \frac{1-\ln 2}{2}\tag{68}$$

and

$$f_1 = f_{hI}(1) = 0.\tag{69}$$

The result (17) together with the expressions (67)-(69) gives the required result (65). □

Result 4.6. *The following bound holds*

$$P_e \le \frac{1}{2}\left[1 - \frac{4}{2\ln 2 - 1}\overline{D}_{I\Delta}(C_1, C_2)\right],\tag{70}$$

where

$$\overline{D}_{I\Delta}(C_1, C_2) = E_X\left\{f_{I\Delta}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\}$$

with

$$f_{I\Delta}^*(x) = \frac{1}{2}\left[\ln 2 + x\ln x + (1-x)\ln(1-x)\right] - (2x-1)^2,\ \forall x \in (0,1).$$

Proof. Let us consider

$$\overline{D}_{I\Delta}(C_1, C_2) = E_X\left\{f_{I\Delta}^*\left(\frac{P(C_1|x)}{P(C_2|x)}\right)P(C_1|x)\right\},$$

where

$$f_{I\Delta}^*(x) = x\, f_{I\Delta}\left(\frac{1-x}{x}\right)$$

$$= \frac{1}{2}\left[\ln 2 + x\ln x + (1-x)\ln(1-x)\right] - (2x-1)^2$$

with

$$f_{I\Delta}(x) = \frac{x}{2}\ln x + \frac{x+1}{2}\ln\left(\frac{2}{x+1}\right) - \frac{(x-1)^2}{4(x+1)}.$$

Now we shall prove the convexity of the function $f_{I\Delta}(x)$. We have

$$f_{I\Delta}'(x) = \frac{1}{2}\ln x + \frac{1}{2}\ln\left(\frac{2}{x+1}\right) - \frac{x-1}{2(x+1)} + \frac{(x-1)^2}{4(x+1)^2}$$

and

$$f_{I\Delta}''(x) = \frac{(x-1)^2}{2x(x+1)^3} \ge 0,\ \forall x > 0.\tag{71}$$

Since, $f_{I\Delta}''(x) \ge 0,\ \forall x > 0$ proving the convexity of the function $f_{I\Delta}(x),\ \forall x > 0$.

We have

$$f^*_{I\Delta}(x) = f^*_{I\Delta}(1-x), \ x \in (0,1), \tag{72}$$

$$f_{(I\Delta)_\infty} = \lim_{x \to \infty} \frac{f_{I\Delta}(x)}{x} = \frac{1}{2} \ln 2 - \frac{1}{4} = \frac{2\ln 2 - 1}{4} \tag{73}$$

and

$$f_1 = f_{I\Delta}(1) = 0. \tag{74}$$

The result (17) together with the expressions (72)-(74) gives the required result (70). $\quad\square$

4.1 Comparison Results

(i) In view of (39) and (71) we have

$$P_e \leq \frac{1}{2}\left[1 - \frac{8}{3}\overline{D}_{h\Delta}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - 4\overline{D}_{I\Delta}(C_1, C_2)\right]. \tag{75}$$

Since $2\ln 2 - 1 > 0$, then from (70) and (75) we have

$$P_e \leq \frac{1}{2}\left[1 - \frac{4}{2\ln 2 - 1}\overline{D}_{I\Delta}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - 4\overline{D}_{I\Delta}(C_1, C_2)\right]. \tag{76}$$

Thus, we observe from (75) that the bound obtained in (70) is sharper than the one obtained applying directly the inequality (39).

(ii) In view of (39) and (65), we have

$$P_e \leq \frac{1}{2}\left[1 - \frac{2}{1 - \ln 2}\overline{D}_{hI}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - \frac{8}{15(1 - \ln 2)}D_{d\Delta}(P||Q)\right]. \tag{77}$$

Since $3 + 8\sqrt{2} - 15\ln 2 > 0$, then from (50) and (77), we have

$$P_e \leq \frac{1}{2}\left[1 - \frac{2}{3 - 2\sqrt{2}}\overline{D}_{dh}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - \frac{8}{15(1 - \ln 2)}D_{d\Delta}(P||Q)\right]. \tag{78}$$

Again, we observe from (78) that the bound obtained in (50) is sharper than the one obtained directly applying the inequality (39).

(iii) In view of (32) and (40), we have

$$P_e \leq \frac{1}{2}\left[1 - \frac{4}{7 - 4\sqrt{2}}\overline{D}_{d\Delta}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - \frac{4}{7 - 4\sqrt{2}}\overline{D}_{h\Delta}(C_1, C_2))\right]. \tag{79}$$

Since $7 - 4\sqrt{2} > 1$, then from (60) and (79) we have

$$P_e \leq \frac{1}{2}\left[1 - 4\overline{D}_{h\Delta}(C_1, C_2)\right] \leq \frac{1}{2}\left[1 - \frac{4}{7 - 4\sqrt{2}}\overline{D}_{h\Delta}(C_1, C_2))\right]. \tag{80}$$

Here also, we observe from (80) that making different comparisons, i.e., instead of using (39) if we use (32) still the result obtained in (60) is sharper. Using the similar arguments we can also make comparisons for the other bounds.

References

[1] E.F. BECKENBACH and R. BELLMAN, *Inequalities.* New York: SpringerVerlag, 1971.

[2] D.E. BOEKEE and J.C. VAN DER LUBBE, Some Aspects of Error Bounds in Feature Selection, *Pattern Recognition*, **11**(1979), 353-360.

[3] J. BURBEA and C.R. RAO, On the Convexity of Some Divergence Measures Based on Entropy Functions, *IEEE Trans. on Inform. Theory*, **IT-28**(1982), 489-495.

[4] I. CSISZÁR, Information Type Measures of Differences of Probability Distribution and Indirect Observations, *Studia Math. Hungarica*, **2**(1967), 299-318.

[5] H. HELLINGER, Neue Begründung der Theorie der quadratischen Formen von unendlichen vielen Veränderlichen, *J. Reine Aug. Math.,* **136**(1909), 210-271.

[6] H. JEFFREYS, An Invariant Form for the Prior Probability in Estimation Problems, *Proc. Roy. Soc. Lon.*, **Ser. A**, **186**(1946), 453-461.

[7] T. KAILATH, The divergence and Bhatttacharyya distance in signal selection, *IEEE Trans. Comm. Tech.*, **COM-15**(1967), 52-60.

[8] S. KULLBACK and R.A. LEIBLER, On Information and Sufficiency, *Ann. Math. Statist.*, **22**(1951), 79-86.

[9] R. SIBSON, Information Radius, *Z. Wahrs. und verw Geb.,* **14**(1969), 149-160.

[10] I.J. TANEJA, On Generalized Information Measures and Their Applications, Chapter in: *Advances in Electronics and Electron Physics*, Ed. P.W. Hawkes, Academic Press, **76**(1989), 327-413.

[11] I.J. TANEJA, New Developments in Generalized Information Measures, Chapter in: *Advances in Imaging and Electron Physics*, Ed. P.W. Hawkes, **91**(1995), 37-136.

[12] I.J. TANEJA and P. KUMAR, Relative Information of Type s, Csiszár $f-$Divergence, and Information Inequalities, *Information Sciences*, **166**(1-4)(2004), 105-125.

[13] I.J. TANEJA , On Symmetric and Nonsymmeric Divergence Measures and Their Generalizations, Chapter in: Advances in Imaging and Electron Physics, **138**(2005), 177-250.

[14] I.J. TANEJA, Refinement of Inequalities among Means, *Journal of Combinatorics, Information and Systems Sciences*, (special issue in honor of Prof. B.D. Sharma) Edited by Prof. Sat Gupta, Vol. 31(2006), 357-378.

[15] I.J. TANEJA, Refinement Inequalities Among Symmetric Divergence Measures, *The Australian Journal of Mathematical Analysis and Applications*, **2**(1)(2005), Art. 8, pp. 1-23.

[16] I.J. TANEJA, Symmetric Divergence Measures and Refinement Inequalities, under preparation.

[17] G.T. TOUSSAINT, On the Divergence between two Distributions and the Probability of Misclassification of Several Decision Rules, *Proc. Second International Joint Conf. on Pattern Recog.*, Copenhagen, Denmark, August, 1-8, 1974.

[18] G.T. TOUSSAINT, Probability of Error, Expected Divergence, and the Affinity of Several Distributions, *IEEE Trans. on Syst. Man and Cybenetics*, **SMC-8**(1978), 482-485.

[19] I. VAJDA, On the $f-$divergence and singularity of probability measures, *Periodica Math. Hunger*, **2**(1972), 223-234.